

Monday 31.01 13h30 – 15h00 Yoav Zemmel

Stochastic approximations for discrete optimal transport

Optimal transport is now a popular tool in statistics, machine learning, and data science. A major challenge in applying optimal transport to large-scale problems is its excessive computational cost. We propose a simple resampling scheme for fast randomised approximate computation of optimal transport distances on finite spaces. This scheme operates on a random subset of the full data and can use any exact algorithm as a black-box back-end, including state-of-the-art solvers and entropically penalised versions. We give non-asymptotic bounds for the expected approximation error. Remarkably, in many important instances such as images (2D-histograms), the bounds are independent of the size of the full problem. Our resampling scheme can also be employed for the barycentre problem, namely computing Fréchet means with respect to the optimal transport metric. We present numerical experiments demonstrating very good approximations could be obtained while decreasing the computation time by several orders of magnitude.

Further recent results on the limiting distributions in discrete optimal transport will also be briefly mentioned.